Gene expression profiles of small-cell lung cancers: Molecular signatures of lung cancer

MASAYA TANIWAKI[1,2], YATARO DAIGO[1], NOBUHISA ISHIKAWA[1], ATSUSHI TAKANO[1], TATSUHIKO TSUNODA[5], WATARU YASUI[3], KOUKI INAI[4], NOBUOKI KOHNO[2] and YUSUKE NAKAMURA[1]

[1]Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo 108-8639; Departments of [2]Molecular and Internal Medicine, [3]Molecular Pathology and [4]Pathology, Graduate School of Biomedical Sciences, Hiroshima University, Hiroshima 734-8551; [5]Laboratory for Medical Informatics, SNP Research Center, RIKEN (Institute of Physical and Chemical Research), Kanagawa 230-0045, Japan

Received March 20, 2006; Accepted May 4, 2006

Abstract. To characterize the molecular mechanisms involved in the carcinogenesis and progression of small-cell lung cancer (SCLC) and identify molecules to be applied as novel diagnostic markers and/or for development of molecular-targeted drugs, we applied cDNA microarray profile analysis coupled with purification of cancer cells by laser-microbeam microdissection (LMM). Expression profiles of 32,256 genes in 15 SCLCs identified 252 genes that were commonly up-regulated and 851 transcripts that were down-regulated in SCLC cells compared with non-cancerous lung tissue cells. An unsupervised clustering algorithm applied to the expression data easily distinguished SCLC from the other major histological type of non-small cell lung cancer (NSCLC) and identified 475 genes that may represent distinct molecular features of each of the two histological types. In particular, SCLC was characterized by altered expression of genes related to neuroendocrine cell differentiation and/or growth such as *ASCL1, NRCAM*, and *INSM1*. We also identified 68 genes that were abundantly expressed both in advanced SCLCs and advanced adenocarcinomas (ADCs), both of which had been obtained from patients with extensive chemotherapy treatment. Some of them are known to be transcription factors and/or gene expression regulators such as *TAF5L, TFCP2L4, PHF20, LMO4, TCF20, RFX2*, and *DKFZp547I048* as well as those encoding nucleotide-binding proteins such as *C9orf76, EHD3*, and *GIMAP4*. Our data provide valuable information for better understanding of lung carcinogenesis and chemoresistance.

Correspondence to: Dr Yusuke Nakamura, Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan
E-mail: yusuke@ims.u-tokyo.ac.jp

Key words: cancer-testis antigen, tumor marker, therapeutic target, expression profile, small-cell lung cancer

Introduction

A number of genetic alterations associated with development and progression of lung cancer, one of the tumors showing the worst prognosis, have been reported, but its precise molecular mechanisms remain unclear (1). Two major histologically-distinct types of lung cancer, non-small cell lung cancer (NSCLC) and small-cell lung cancer (SCLC) have different pathophysiological and clinical features that suggest differences in the mechanisms of their carcinogenesis. SCLC accounts for 15-20% of all lung cancers (2,3) and is categorized as neuroendocrine tumors of the lung with certain morphologic, ultrastructural, and immunohistochemical characteristics. Some paraneoplastic syndromes such as inappropriate secretion of antidiuretic hormone, ectopic Cushing's syndrome, and the Lambert-Eaton myasthenic syndrome (LEMS) are known to be associated with SCLC, however, the detailed molecular characteristics of neuroendocrine tumors are still not well understood. Although patients with SCLC respond favorably to 1st line multi-agent chemotherapy, they often relapse in a short time. Hence, only 20% of patients with limited-stage disease (LD) can be cured with combined modality therapy and <5% of those with extensive-disease (ED) can achieve 5-year survival after the initial diagnosis (4,5). Therefore, new therapeutic strategies such as molecular-targeted agents are eagerly awaited.

The genome-wide cDNA microarray analysis enabled us to obtain comprehensive gene expression profiles related to detailed phenotypic and biological information in cancer cells (6-8). This approach is also useful to identify unknown molecules involved in the pathways of lung carcinogenesis (9-11). We, in fact, have identified multiple applicable targets for development of novel anti-cancer drugs and/or diagnostic markers (12-17).

Through gene-expression profile analysis of 15 SCLCs coupled with purification of cancer cell population by laser-microbeam microdissection (LMM) on a cDNA microarray consisting of 32,256 cDNAs, we identified a number of genes that were expressed differently between the two most common histological types of lung cancer, NSCLC and SCLC. We report important information regarding the mechanisms of

lung carcinogenesis and chemoresistance as well as the discovery of potential targets for development of diagnostic markers and signal-suppressing therapeutic strategies for lung cancer treatment.

Materials and methods

Patients and tissue samples. Advanced SCLC tissue samples were obtained with informed consent from post-mortem materials (15 individuals) at Hiroshima University (Hiroshima, Japan). Individual institutional ethical committees approved the use of all clinical materials. Patients' clinical information was obtained from medical records. All samples were immediately frozen and embedded in TissueTek OCT medium (Sakura, Tokyo, Japan) and stored at -80˚C until use for microarray analysis.

Laser-microbeam microdissection, extraction of RNA and T7-based RNA amplification. Cancer cells were selectively collected from the preserved samples using laser-microbeam microdissection (10,11). The quality of RNAs extracted from the residual tissue of each case was checked by electrophoresis in the degenerative agarose gel and ensured by a presence of clear ribosomal RNA bands. Extraction of total RNA and T7-based amplification were performed as described previously (10,11). As a control probe, normal human lung poly(A) RNA (BD Biosciences Clontech, Palo Alto, CA) was amplified in the same manner; 2.5 μg each of amplified RNAs (aRNAs) from each cancerous tissue and that from the control were reversely transcribed in the presence of Cy5-dCTP and Cy3-dCTP, respectively.

cDNA microarrays. Our genome-wide cDNA microarray system containing 32,256 cDNAs that were selected on the basis of the information in the UniGene database (build #188) of the National Center for Biotechnology Information (NCBI) was used for this analysis. Fabrication of the microarray, hybridization, washing, and detection of signal intensities were described previously (10,18).

Data analysis. Signal intensities of Cy3 and Cy5 from the 32,256 spots were quantified and analyzed by substituting backgrounds, using ArrayVision software (Imaging Research Inc., Ontario, Canada). Subsequently, the fluorescent intensities of Cy5 (tumor) and Cy3 (control) for each target spot were adjusted so that the mean Cy5/Cy3 ratio of 52 housekeeping genes on the array was equal to one. Because data derived from low signal intensities are less reliable, we determined a cut-off value on each slide as described previously (10) and excluded genes from further analysis when both Cy3 and Cy5 dyes yielded signal intensities lower than the cut-off. For other genes, we calculated the Cy5/Cy3 ratio using the raw data of each sample.

Cluster analysis of SCLC and NSCLCs. We applied a hierarchical clustering method to both genes and tumors. To obtain reproducible clusters for classification of the 15 SCLC, and an independent set of 62 NSCLC (20 early-stage ADC, 15 early-stage SCC, and 27 advanced ADC) samples analyzed previously using a cDNA microarray containing a subset (27,648 genes) of 32,256 genes on our present microarray-system (data from refs. 9 and 11, and our unpublished data for the expression of 4608 genes in the same set of 35 early-stage NSCLC), we selected genes from them for which valid data were obtained in 80% of the experiments, and whose expression ratios varied by standard deviations of >1.7. The analysis was performed using web-available software ('Cluster' and 'TreeView') written by M. Eisen (http://rana.lbl.gov/ index.htm). Before applying the clustering algorithm, we log-transformed the fluorescence ratio for each spot and then median-centered the data for each sample to remove experimental biases.

Semi-quantitative RT-PCR. We selected highly up-regulated genes and examined their expression levels by means of semi-quantitative RT-PCR experiments as previously described (16). aRNA (3 μg) from each sample was reversely transcribed to single-stranded cDNAs using random primer (Roche) and Superscript II (Invitrogen). Semi-quantitative RT-PCR experiments were carried out with the following sets of synthesized primers specific to the 10 representative genes that were up-regulated in SCLCs or with ß-actin (*ACTB*)-specific primers as an internal control: K562 cell-derived leucine-zipper-like protein 1 (*KLP1*), 5'-CCGTCAGCAGTG TGAAGTCT-3' and 5'-CCTCCTAAGCAGTCAACCTTGT-3'; BCL2-associated athanogene 5 (*BAG5*), 5'-ATCTGGTTT TTAAGGGTCTGAGC-3' and 5'-GCAAGCGTAAGAGAC TGGTTTTA-3'; solute carrier family 4, sodium bicarbonate cotransporter, member 5 (*SLC4A5*), 5'-CTGTCAGGGTCAT AGTAGGCATT-3' and 5'-CCAAAGTCAAACTCCCATT CAT-3'; hypothetical protein FLJ13848 (*FLJ13848*), 5'-AA AGAGGAACACACTGGGTGTAA-3' and 5'-AGGAGCC TAGAGAAGCAATCATC-3'; brain-specific angiogenesis inhibitor 3 (*BAI3*), 5'-TTGCTTCCTAATCCCTTTGGTC-3' and 5'-TAAGCTGCATCTTGATGCCTTC-3'; vitelliform macular dystrophy 2-like 3 (*VMD2L3*), 5'-CCAGTTACTGT GTCTATCGGGTC-3' and 5'-AGCCATATGTAGTCAAGT GCCAT-3'; cell division cycle 25C (*CDC25C*), 5'-CAGATG CTGGAGGAAGATTCTAA-3' and 5'-AAAGAAAGAGGG GGAAACAAAG-3'; secretory carrier membrane protein 5 (*SCAMP5*), 5'-AGAACTTTGGCTCCCTTTCC-3' and 5'-TGCATAGTTGCCTGGAGATG-3'; ectodermal-neural cortex (*ENC1*), 5'-GTCCATGCCATGAATGAGTG-3' and 5'-CTC TTGGCAGATTTGCATCA-3'; CDC20 cell division cycle 20 homolog (*CDC20*), 5'-CCTCTGGTCTCCCCATTACA-3' and 5'-CTGAGGTGATGGGTTGGTCT-3'; ß-actin (*ACTB*), 5'-GAGGTGATAGCATTGCTTTCG-3' and 5'-CAAGTCA GTGTACAGGTAAGC-3'. PCR reactions were optimized for the number of cycles to ensure the PCR product is within the linear phase of amplification.

Immunohistochemical analysis. To confirm the differential protein expression of 2 candidate markers (SCAMP5 and CDC20), which were highly up-regulated in SCLC, we stained clinical tissue sections using ENVISION+ Kit/HRP (DakoCytomation). Briefly, after endogenous peroxidase and protein blocking reactions, anti-human SCAMP5 polyclonal antibody (Medical & Biological Laboratories, Aichi, Japan) or anti-human CDC20 monoclonal antibody (Santa Cruz Biotechnology, CA) was added, and then HRP-labeled

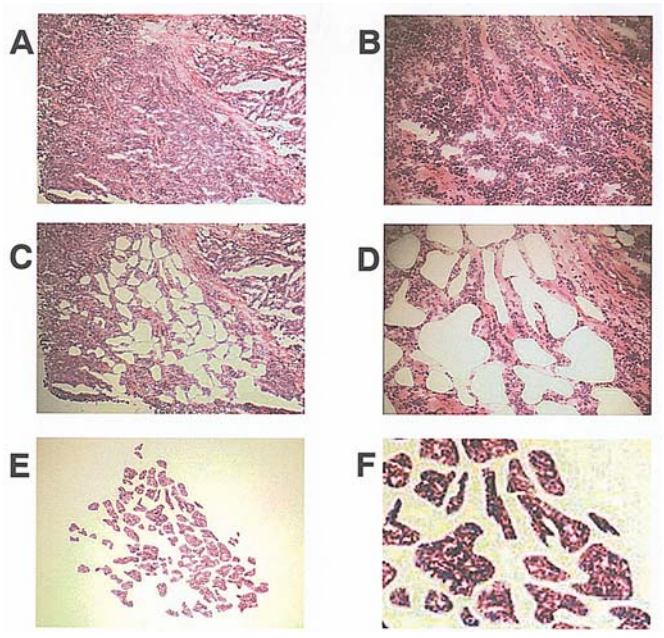

Figure 1. Images illustrating laser-microbeam microdissection (LMM) of two representative SCLCs (each case corresponds to either left or right panel). The upper two panels (A and B) show the samples before dissection; the middle (C and D), the same sections after microdissection (H&E stain). The microdissected cancer cells captured on the collecting cap were also shown in the bottom panels (E and F).

anti-rabbit or anti-mouse IgG as the secondary antibody. Substrate-chromogen was then added and the specimens were counterstained with hematoxylin.

Results

Identification of genes commonly up- and down-regulated in SCLCs. To obtain precise gene expression profiles of SCLC cells, we collected pure populations of cancer cells by LMM to minimize contamination of the tumor cells by non-cancerous cells (Fig. 1). We identified up- and down-regulated genes common to SCLC according to the following criteria: i) genes for which we were able to obtain expression data in >80% (at least twelve of the 15 cases) of the cases examined; and ii) genes whose expression ratio was >10.0 or <0.1 in at least 50% of the informative cases. According to these criteria, a total of 252 genes were listed up as commonly up-regulated and 851 genes as commonly down-regulated in SCLC (representative genes are listed in Table I). The up-regulated genes represented a variety of functions including genes associated with cell adhesion and cytoskeleton, signal transduction, cell proliferation, and some kinases (Table I). Some of the transactivated genes included in the list, such as GRP, SKP2 and INSM1 were previously reported by others (19-21).

Validation of selected genes by semi-quantitative RT-PCR and immunohistochemical analyses. To validate the expression data obtained by microarray analysis, we performed semi-quantitative RT-PCR experiments for 10 representative genes, which were indicated to be overexpressed commonly in SCLCs (Fig. 2A). The results of RT-PCR experiments were quite

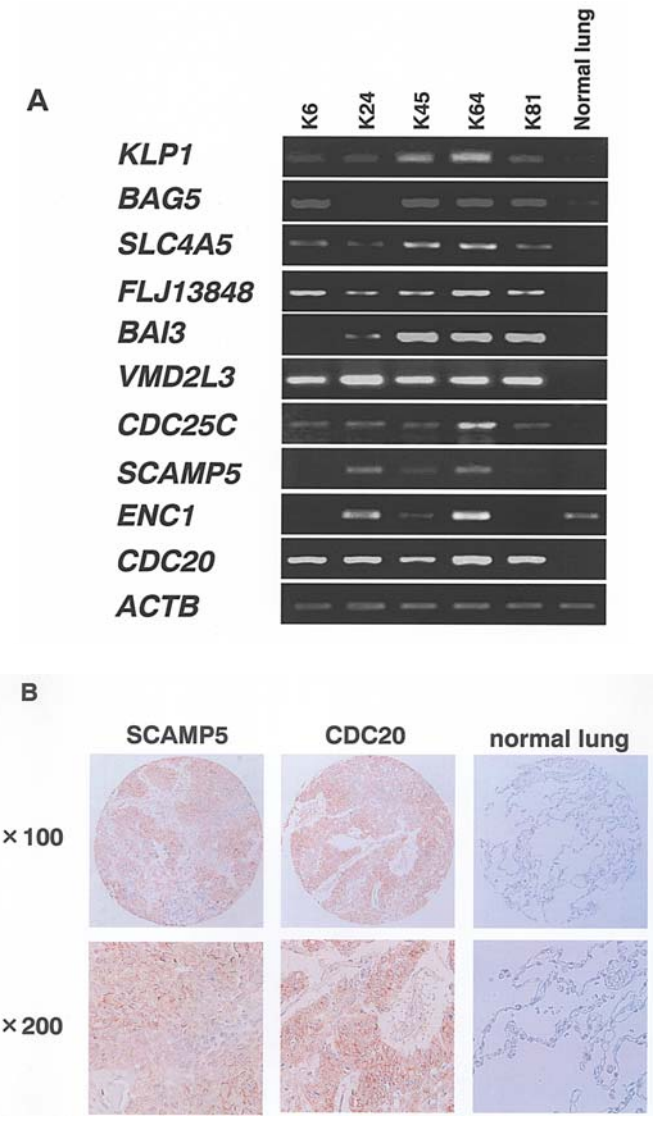

Figure 2. Validation of the representative gene and protein expression by semi-quantitative RT-PCR and immunohistochemical analyses. A, representative image of semi-quantitative RT-PCR analysis of RNAs from the SCLC samples used in this study. The integrity of each cDNA template was controlled through amplification of *ACTB*. B, immunohistochemical staining of representative samples from the SCLCs and normal lung tissue examined, using antibodies for 2 candidate protein markers (SCAMP5 and CDC20) (x100, x200).

concordant with those of the microarray data. To further validate the data at the protein level, we carried out immunohistochemical analysis using the paired tumor and normal tissue sections using antibodies for SCAMP5 or CDC20 (Fig. 2B). Both proteins were confirmed to be expressed abundantly in SCLCs, but were hardly detectable in normal lung.

Cluster analysis of gene expression in lung cancers. We then applied an unsupervised two-dimensional hierarchical clustering algorithm to analyze similarities among samples and among genes, using data obtained from expression profiles of 15 advanced SCLCs as well as 35 early-stage NSCLCs (ADC and SCC) and 27 advanced NSCLCs (ADC) which we previously analyzed and reported (9,11) (Fig. 3A). Since the

Table I. Representative up-regulated genes in SCLC with known function.

GenBank ID	Symbol	Gene name
Cell adhesion and cytoskeleton		
AY714129	*CELSR3*	Cadherin, EGF LAG seven-pass G-type receptor 3 (flamingo homolog, *Drosophila*)
BX537667	*FARP1*	FERM, RhoGEF (ARHGEF) and pleckstrin domain protein 1 (chondrocyte-derived)
S78296	*INA*	Internexin neuronal intermediate filament protein, α
Signal transduction		
BI496673	*BAI3*	Brain-specific angiogenesis inhibitor 3
BC034227	*D4S234E*	DNA segment on chromosome 4 (unique) 234 expressed sequence
NM_000555	*DCX*	Doublecortex; lissencephaly, X-linked (doublecortin)
R20639	*DPYSL5*	Dihydropyrimidinase-like 5
BC014476	*GKAP1*	G kinase anchoring protein 1
Cell proliferation		
BC010044	*CDC20*	CDC20 cell division cycle 20 homolog (*S. cerevisiae*)
NM_001790	*CDC25C*	Cell division cycle 25C
NM_031966	*CCNB1*	Cyclin B1
AF053306	*BUB1B*	BUB1 budding uninhibited by benzimidazoles 1 homolog ß (yeast)
AF260237	*HES6*	Hairy and enhancer of split 6 (*Drosophila*)
BC000356	*MAD2L1*	MAD2 mitotic arrest deficient-like 1 (yeast)
Enzymatic activities		
M76180	*DDC*	Dopa decarboxylase (aromatic L-amino acid decarboxylase)
NM_020546	*ADCY2*	Adenylate cyclase 2 (brain)
X60673	*AK3*	Adenylate kinase 3-like 1
AF055015	*EYA2*	Eyes absent homolog 2 (*Drosophila*)

62 NSCLC samples had been analyzed for a subset (27,648 genes) of the 32,256 genes on our present microarray-system, we analyzed the information of a subset of the 27,648 genes for which valid values could be obtained in >80% of the cases examined. We also excluded genes with observed standard deviations of <1.7. The 475 genes that passed through this cut-off filter were analyzed further.

In the sample axis (horizontal) in Fig. 3A, 81 samples (four cases were examined in duplicate to validate the reproducibility and reliability of our experimental procedure), 77 cases were clustered into two major groups on the basis of their expression profiles. The dendrogram shown at the top of Fig. 3 represents similarities in expression patterns among individual cases; the shorter the branches are, the greater the similarities are. The four duplicated cases (no. 13, 20, K91, and LC12) that were labelled and hybridized in independent experiments were clustered most closely within the same group (Fig. 3B). The identical genes spotted on different positions on the slide glasses were also clustered into adjacent rows (Fig. 3B). These results supported the high reproducibility and reliability of our experimental procedures. Of the 77 cases, 15 SCLC clustered into one major group and 20 early-stage ADC and 15 SCC as well as 27 advanced ADC clustered into individual groups. Clearly, SCLC and NSCLC appeared to have different gene expression profiles that could reflect differences in the etiological and clinicopathological natures.

In this analysis, we obtained 34 genes which were expressed abundantly in SCLC, some of which revealed characteristics of certain neuronal functions such as neurogenesis and neuroprotection (Cluster-1 in Fig. 3A and B; Table IIA; i.e. *DPYSL2, ADNP* etc).

Identification of genes related to chemoresistance. Since chemoresistance is a major obstacle for cancer treatment, identification of genes commonly up-regulated in cancer cells obtained from patients who had failed certain chemotherapy is an effective approach to understand the mechanism of chemoresistance and develop a novel cancer therapy that overcomes this problem. We obtained 68 genes expressed abundantly both in advanced SCLCs and advanced ADCs (Cluster-2 in Fig. 3A and C; Table IIB), both of which were obtained from patients who had extensive chemotherapy treatments (although the chemotherapy protocols provided to these patients were not the same). Some of them are known to be transcription factors and/or gene expression regulators such as *TAF5L, TFCP2L4, PHF20, LMO4, TCF20, RFX2*,

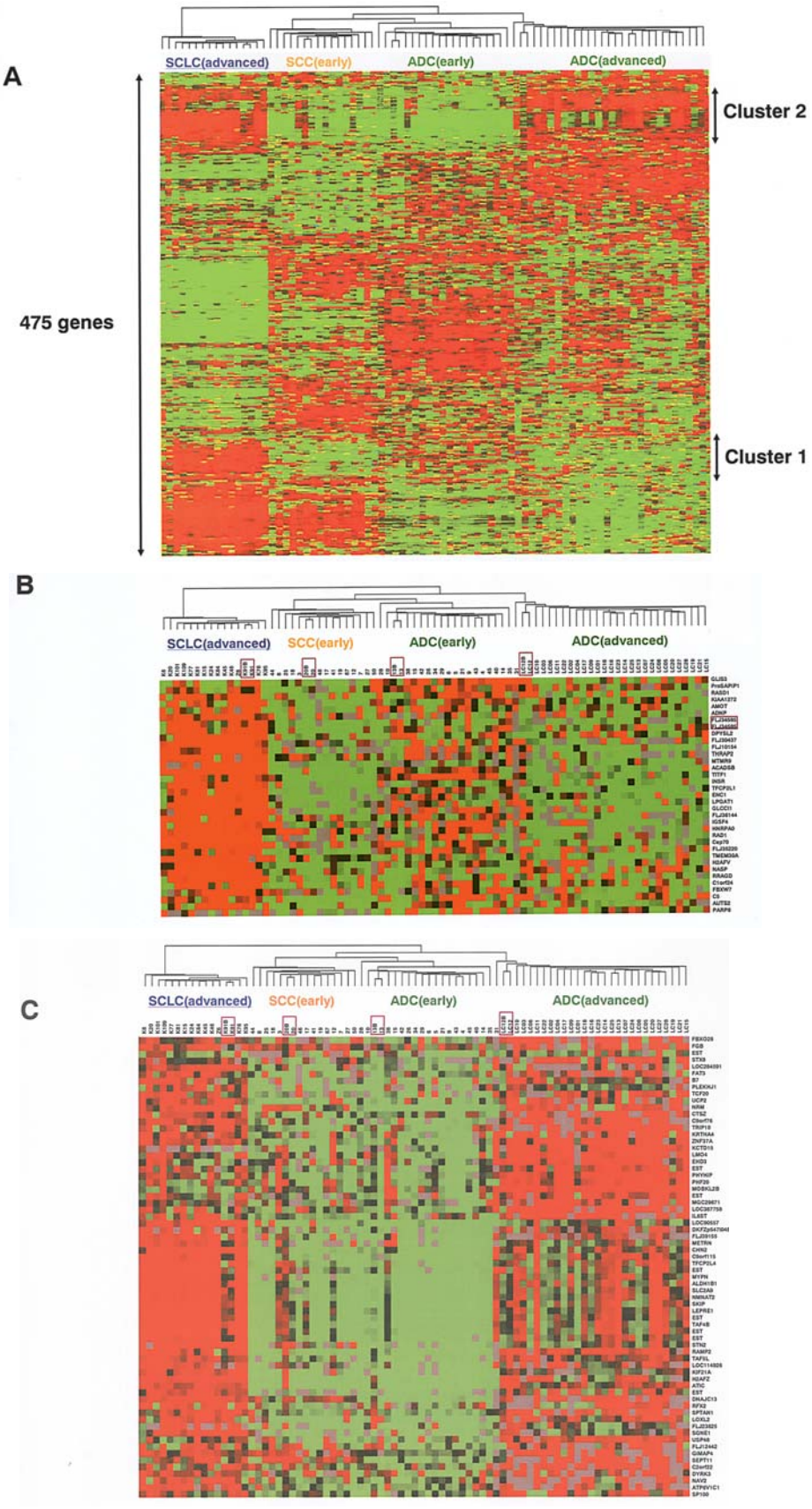

Figure 3. Dendrogram of two-dimensional hierarchical clustering analysis of genes across samples from 77 lung cancer cases. Red and green indicate transcript levels respectively above and below the median for the respective gene across all samples. Black, unchanged expression; gray, no detectable expression. In the horizontal axis representing 77 lung cancers, 15 advanced SCLCs, 35 early-stage NSCLCs (20 ADCs and 15 SCCs) and 27 advanced ADCs were separated in four trunks. In the vertical axis the 475 genes were clustered in different branches according to similarities in relative expression ratio (A). Cluster-1 includes genes which expressed more abundantly in SCLCs than in NSCLCs. The four duplicated cases (no. 13, 20, K91, and LC12) that were labelled and hybridized in independent experiments were clustered most closely within the same group. Identical genes spotted on different positions on the slide glasses were also clustered into adjacent rows (orange square) (B). Cluster-2 includes genes which commonly expressed in advanced SCLCs and NSCLCs, both of which had been treated with chemotherapy (C).

Table II. A, list of 34 genes expressed abundantly in SCLCs.

No.	GenBank ID	Symbol	Gene name
1	AB209404	*GLIS3*	GLIS family zinc finger 3
2	AB011124	*ProSAPiP1*	ProSAPiP1 protein
3	BC042688	*RASD1*	RAS, dexamethasone-induced 1
4	AK022881	*KIAA1272*	Chromosome 20 open reading frame 74
5	NM_133265	*AMOT*	Angiomotin
6	CA503163	*ADNP*	Activity-dependent neuroprotector
7	AA058578	*FLJ34585*	CDNA FLJ34585 fis, clone KIDNE2008758
8	BX647115	*DPYSL2*	Dihydropyrimidinase-like 2
9	AK054999	*FLJ30437*	CDNA FLJ30437 fis, clone BRACE2009045
10	BQ016211	*FLJ10154*	Hypothetical protein FLJ10154
11	AA418594	*THRAP2*	Thyroid hormone receptor associated protein 2
12	NM_015458	*MTMR9*	Myotubularin related protein 9
13	NM_001609	*ACADSB*	Acyl-Coenzyme A dehydrogenase, short/branched chain
14	U33749	*TITF1*	Thyroid transcription factor 1
15	AL365454	*INSR*	Insulin receptor
16	AI928242	*TFCP2L1*	Transcription factor CP2-like 1
17	AF059611	*ENC1*	Ectodermal-neural cortex (with BTB-like domain)
18	AA921341	*LPGAT1*	Lysophosphatidylglycerol acyltransferase 1
19	AA602499	*GLCCI1*	Glucocorticoid induced transcript 1
20	AK124953	*FLJ36144*	Similar to hypothetical protein FLJ36144
21	R42757	*IGSF4*	Immunoglobulin superfamily, member 4
22	CR596214	*HNRPA0*	Heterogeneous nuclear ribonucleoprotein A0
23	AK096960	*RAD1*	RAD1 homolog (S. pombe)
24	AI341170	*Cep70*	P10-binding protein
25	AK096344	*FLJ35220*	Hypothetical protein FLJ35220
26	AL832815	*TMEM30A*	Transmembrane protein 30A
27	AL110212	*H2AFV*	H2A histone family, member V
28	NM_172164	*NASP*	Nuclear autoantigenic sperm protein (histone-binding)
29	N29574	*RRAGD*	Ras-related GTP binding D
30	AL137572	*C1orf24*	Chromosome 1 open reading frame 24
31	NM_033632	*FBXW7*	F-box and WD-40 domain protein 7 (archipelago homolog, *Drosophila*)
32	AA788924	*C5*	Complement component 5
33	AF326917	*AUTS2*	Autism susceptibility candidate 2
34	BQ002875	*PARP8*	Poly(ADP-ribose) polymerase family, member 8

B, list of 68 genes expressed abundantly both in advanced SCLCs and advanced ADCs.

No.	GenBank ID	Symbol	Gene name
1	AB007952	*FBXO28*	F-box protein 28
2	NM_005141	*FGB*	Fibrinogen ß chain
3	AA830326	*EST*	
4	AA677491	*STX8*	Syntaxin 8
5	AK091100	*LOC284591*	Hypothetical protein LOC284591
6	AA464854	*FAT3*	FAT tumor suppressor homolog 3 (*Drosophila*)
7	BC029858	*B7*	B7 gene
8	CA306079	*PLEKHJ1*	Pleckstrin homology domain containing, family J member 1
9	AA634326	*TCF20*	Transcription factor 20 (AR1)
10	AK025742	*UCP2*	Uncoupling protein 2 (mitochondrial, proton carrier)
11	AK075509	*NRM*	Nurim (nuclear envelope membrane protein)
12	NM_001336	*CTSZ*	Cathepsin Z
13	BC039999	*C9orf76*	Chromosome 9 open reading frame 76
14	AF502289	*TRIP10*	Thyroid hormone receptor interactor 10
15	BC041070	*KRTHA4*	Keratin, hair, acidic, 4

Table II. B, continued.

No.	GenBank ID	Symbol	Gene name
16	NM_001007094	*ZNF37A*	Zinc finger protein 37a (KOX 21)
17	AA868706	*KCTD15*	Potassium channel tetramerisation domain containing 15
18	CV424097	*LMO4*	LIM domain only 4
19	AF214736	*EHD3*	EH-domain containing 3
20	AA757392	*EST*	
21	D87463	*PHYHIP*	Phytanoyl-CoA hydroxylase interacting protein
22	BM916826	*PHF20*	PHD finger protein 20
23	H12117	*MOBKL2B*	MOB1, Mps One Binder kinase activator-like 2B (yeast)
24	R32836	*EST*	
25	AA563634	*MGC29671*	Hypothetical protein MGC29671
26	NM_203371	*LOC387758*	Similar to RIKEN cDNA 1110018M03
27	NM_002184	*IL6ST*	Interleukin 6 signal transducer (gp130, oncostatin M receptor)
28	AK097664	*LOC90557*	Hypothetical protein BC016861
29	AA813719	*DKFZp547I048*	Chromosome 1 open reading frame 173
30	NM_182798	*FLJ39155*	Hypothetical protein FLJ39155
31	AK057053	*METRN*	Meteorin, glial cell differentiation regulator
32	H11638	*CHN2*	Chimerin (chimaerin) 2
33	N93264	*C9orf115*	Chromosome 9 open reading frame 115
34	BC036890	*TFCP2L4*	Grainyhead-like 3 (*Drosophila*)
35	BX109199	*EST*	
36	AL834247	*MYPN*	Myopalladin
37	NM_000692	*ALDH1B1*	Aldehyde dehydrogenase 1 family, member B1
38	R49124	*SLC2A9*	Solute carrier family 2 (facilitated glucose transporter), member 9
39	AA828735	*NMNAT2*	Nicotinamide nucleotide adenylyltransferase 2
40	CR749297	*SKIP*	SPHK1 (sphingosine kinase type 1) interacting protein
41	AF097431	*LEPRE1*	Leucine proline-enriched proteoglycan (leprecan) 1
42	BG209407	*EST*	Transcribed locus
43	AI347994	*TAF4B*	TAF4b RNA polymerase II, TATA box binding protein (TBP)-associated factor, 105 kDa
44	BU628989	*EST*	
45	AA429665	*EST*	
46	BX648249	*STN2*	Stonin 2
47	N93656	*RAMP2*	Receptor (calcitonin) activity modifying protein 2
48	NM_014409	*TAF5L*	TAF5-like RNA polymerase II, p300/CBP-associated factor (PCAF)-associated factor, 65 kDa
49	W52081	*LOC114926*	Hypothetical protein BC013035
50	AF450487	*KIF21A*	Kinesin family member 21A
51	BM472056	*H2AFZ*	H2A histone family, member Z
52	CR606023	*ATIC*	5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase/IMP cyclohydrolase
53	H05226	*EST*	
54	AB014578	*DNAJC13*	DnaJ (Hsp40) homolog, subfamily C, member 13
55	AI288717	*RFX2*	Regulatory factor X, 2 (influences HLA class II expression)
56	BC053521	*SPTAN1*	Spectrin, α, non-erythrocytic 1 (α-fodrin)
57	U89942	*LOXL2*	Lysyl oxidase-like 2
58	BC035561	*FLJ23825*	Hypothetical protein FLJ23825
59	BC093053	*SGNE1*	Secretory granule, neuroendocrine protein 1 (7B2 protein)
60	NM_032236	*USP48*	Ubiquitin specific protease 48
61	AK023995	*FLJ12442*	Hypothetical protein FLJ12442
62	NM_018326	*GIMAP4*	GTPase, IMAP family member 4
63	NM_018243	*SEPT11*	Septin 11
64	AA195424	*C2orf22*	PQ loop repeat containing 3
65	Y12735	*DYRK3*	Dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 3
66	NM_182964	*NAV2*	Neuron navigator 2
67	NM_001695	*ATP6V1C1*	ATPase, H^+ transporting, lysosomal 42 kDa, V1 subunit C, isoform 1
68	U36501	*SP100*	Nuclear antigen Sp100

and *DKFZp547I048*. Furthermore, some genes encoding nucleotide-binding proteins such as *C9orf76, EHD3*, and *GIMAP4* were also found in the list.

Discussion

Chemotherapy remains as the essential component for treatment of patients with SCLC, regardless of their stage (either LD or ED) or performance status. In LD, a combination of chemotherapy with radiation therapy improves patient prognosis better than chemotherapy alone (22,23). SCLC is usually sensitive to chemotherapy and radiotherapy in first-line treatment, but the period of good response is often very limited and most of the patients ultimately relapse with treatment-resistant disease. Hence, the final outcome of SCLC patients is poor with an overall 5-year survival rate of <10%. Therefore, it is urgently required to develop novel diagnostic tools for detection of cancer at a much earlier stage and to develop molecular-targeted therapies such as small-molecular compound and immunotherapies targeting cancer-specific molecules. Gene expression profile of SCLC is the first step to identify the appropriate candidate molecules for such purposes. To analyze the gene expression profile of SCLC, we used genome-wide cDNA microarray consisting of 32,256 cDNA clones in combination with laser-microdissection technology to represent the expression profile of a highly pure population of SCLC cells. To our knowledge, this is the first report of comprehensive gene expression profiles of pure SCLC cells derived from laser-microdissected clinical tissue samples.

Through the detailed genome-wide expression data of 32,256 genes, we identified 252 transcripts that were commonly up-regulated in SCLCs. These genes encode proteins with a variety of functions that include transmembrane/secretory proteins, and cancer-testis or onco-fetal antigens as well as proteins important in cell adhesion, cytoskeleton structure, signal transduction, and cell proliferation. Some of them should be useful as diagnostic/prognostic markers and probably as therapeutic targets for development of new molecular-targeted agents or immunotherapy for lung-cancer treatment. Tumor-specific transmembrane/secretory proteins should have significant advantages, because they are presented on the cell surface, making them easily accessible as molecular markers and therapeutic targets. Some tumor-specific markers available at present, such as CYFRA or Pro-GRP, are transmembrane/secretory proteins (19,24). An example of rituximab (Rituxan), a chimeric monoclonal antibody against CD20-positive lymphomas, provides proof of the concept that targeting specific cell-surface proteins can provide significant clinical benefits (25). On the other hand, among tumor antigens identified to date, cancer-testis antigens (CTAs) have been recognized as a group of highly attractive targets for cancer vaccine. Although other factors, such as the *in vivo* immunogenicity of the protein are also important and further examination will be necessary, our candidate genes include known CTA such as TSGA14. Further study using this expression profile will doubtlessly enable us to identify novel CTAs that could be a good target for immunotherapy of SCLC.

Chemoresistance is clinically a very important issue that we need to overcome for an improvement in the treatment of patients with advanced or end-stage cancer. Our gene expression profile data obtained from the fifteen autopsy samples as well as advanced ADCs with the clinical history of chemotherapy (Cluster-2 in Fig. 3A and C; Table IIB) were considered to reflect the characteristics of advanced lung cancers with acquired chemoresistance. Unsupervised cluster analysis of these subgroups identified up-regulated genes including *TAF5L, TFCP2L4, PHF20, LMO4, TCF20*, and *RFX2* that were known to have transcription factor activities. Some transcription factors were reported to be associated with acquired chemoresistance. For example, constitutive activation of NF-κB, a transcription factor involved in multiple cellular processes, appears to support cancer cell survival and to reduce the sensitivity against chemotherapeutic drugs (26). On the other hand, some genes in the list, i.e. *C9orf76, EHD3*, and *GIMAP4*, were found to bind to the nucleotide. Since some DNA-binding proteins were known to play a critical role in the DNA-repair process, the genes shown above might also have some functions in DNA repair and contribute to increase in chemoresistance. Further analysis of the genes in this group might be important for the development of novel therapies for chemoresistant tumors.

Neuroendocrine tumors of the lung range from well-differentiated neuroendocrine carcinoma (typical carcinoid) to intermediate grade (atypical carcinoma) or to very aggressive poorly differentiated lesions (large cell neuroendocrine carcinoma (LCNEC) and SCLC). SCLC is generally considered as a major neuroendocrine tumor of the lung, and causes several paraneoplastic neuroendocrine syndromes. These syndromes represent clinically distinct symptoms in SCLC patients. Up-regulated genes included several genes which were related to the neuroendocrine function such as insulinoma-associated 1 (*INSM1*), chromogranin A (parathyroid secretory protein 1; *CHGA*), and achaete-scute complex-like 1 (*Drosophila*; *ASCL1*), further supporting the strong relationship between SCLC and neuroendocrine syndromes at molecular levels. Our gene list might also include a set of genes related to some cancer-related syndromes including cachexia.

In conclusion, our cDNA microarray analysis combined with an LMM system revealed most comprehensive gene expression profiles of SCLC involving up-regulated genes that encode proteins with the function of cell cycle/growth, and signal transduction, or products with unknown function as well as transmembrane/secretory proteins and CTAs. The information could offer a powerful strategy for rapid identification and further evaluation of target molecules for a personalized therapy.

Acknowledgements

This work was supported in part by a 'Research for the Future' Program Grant of The Japan Society for the Promotion of Science (#00L01402) to Y.N.

References

1. Sozzi G: Molecular biology of lung cancer. Eur J Cancer 37 (suppl 7): S63-S73, 2001.
2. Morita T and Sugano H: A statistical analysis of lung cancer registered in the Annual of Pathological Autopsy Cases in Japan between 1958 and 1987, with special reference to the characteristics of lung cancer in Japan. Acta Pathol Jpn 40: 665-675, 1990.

3. Simon GR and Wagner H: Small cell lung cancer. Chest 123 (suppl 1): S259-S271, 2003.
4. Chute JP, Chen T, Feigal E, Simon R and Johnson BE: Twenty years of phase III trials for patients with extensive-stage small-cell lung cancer: perceptible progress. J Clin Oncol 17: 1794-1801, 1999.
5. Sandler AB: Chemotherapy for small cell lung cancer. Semin Oncol 30: 9-25, 2003.
6. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD and Lander ES: Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 286: 531-537, 1999.
7. Pomeroy SL, Tamayo P, Gaasenbeek M, Sturla LM, Angelo M, McLaughlin ME, Kim JY, Goumnerova LC, Black PM, Lau C, Alien JC, Zagzag D, Olson JM, Curran T, Wetmore C, Biegel JA, Poggio T, Mukherjee S, Rifkin R, Califano A, Stolovitzky G, Louis DN, Mesirov JP, Lander ES and Golub TR: Prediction of central nervous system embryonal tumour outcome based on gene expression. Nature 415: 436-442, 2002.
8. Van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R and Friend SH: Gene expression profiling predicts clinical outcome of breast cancer. Nature 415: 530-536, 2002.
9. Kikuchi T, Daigo Y, Katagiri T, Tsunoda T, Okada K, Kakiuchi S, Zembutsu H, Furukawa Y, Kawamura M, Kobayashi K, Imai K and Nakamura Y: Expression profiles of non-small cell lung cancers on cDNA microarrays: identification of genes for prediction of lymph-node metastasis and sensitivity to anti-cancer drugs. Oncogene 22: 2192-2205, 2003.
10. Kakiuchi S, Daigo Y, Tsunoda T, Yano S, Sone S and Nakamura Y: Genome-wide analysis of organ-preferential metastasis of human small cell lung cancer in mice. Mol Cancer Res 1: 485-499, 2003.
11. Kakiuchi S, Daigo Y, Ishikawa N, Furukawa C, Tsunoda T, Yano S, Nakagawa K, Tsuruo T, Kohno N, Fukuoka M, Sone S and Nakamura Y: Prediction of sensitivity of advanced non-small cell lung cancers to gefitinib (Iressa, ZD1839). Hum Mol Genet 13: 3029-3043, 2004.
12. Suzuki C, Daigo Y, Kikuchi T, Katagiri T and Nakamura Y: Identification of COX17 as a therapeutic target for non-small cell lung cancer. Cancer Res 63: 7038-7041, 2003.
13. Suzuki C, Daigo Y, Ishikawa N, Kato T, Hayama S, Ito T, Tsuchiya E and Nakamura Y: ANLN plays a critical role in human lung carcinogenesis through activation of RHOA and by involvement in PI3K/AKT pathway. Cancer Res 65: 11314-11325, 2005.
14. Ishikawa N, Daigo Y, Yasui W, Inai K, Nishimura H, Tsuchiya E, Kohno N and Nakamura Y: ADAM8 as a novel serological and histochemical marker for lung cancer. Clin Cancer Res 10: 8363-8370, 2004.
15. Ishikawa N, Daigo Y, Takano A, Taniwaki M, Kato T, Hayama S, Murakami H, Takeshima Y, Inai K, Nishimura H, Tsuchiya E, Kohno N and Nakamura Y: Increases of amphiregulin and transforming growth factor-alpha in serum as predictors of poor response to gefitinib among patients with advanced non-small cell lung cancers. Cancer Res 65: 9176-9184, 2005.
16. Kato T, Daigo Y, Hayama S, Ishikawa N, Yamabuki T, Ito M, Miyamoto M, Kondo S and Nakamura Y: A novel human tRNA-dihydrouridine synthase involved in pulmonary carcinogenesis. Cancer Res 65: 5638-5646, 2005.
17. Furukawa C, Daigo Y, Ishikawa N, Kato T, Ito T, Tsuchiya E, Sone S and Nakamura Y: PKP3 oncogene as prognostic marker and therapeutic target for lung cancer. Cancer Res 65: 7102-7110, 2005.
18. Ochi K, Daigo Y, Katagiri T, Nagayama S, Tsunoda T, Myoui A, Naka N, Araki N, Kudawara I, Ieguchi M, Toyama Y, Toguchida J, Yoshikawa H and Nakamura Y: Prediction of response to neoadjuvant chemotherapy for osteosarcoma by gene-expression profiles. Int J Oncol 24: 647-655, 2004.
19. Miyake Y, Kodama T and Yamaguchi K: Pro-gastrin-releasing peptide(31-98) is a specific tumor marker in patients with small cell lung carcinoma. Cancer Res 54: 2136-2140, 1994.
20. Yokoi S, Yasui K, Saito-Ohara F, Koshikawa K, Iizasa T, Fujisawa T, Terasaki T, Horii A, Takahashi T, Hirohashi S and Inazawa J: A novel target gene, SKP2, within the 5p13 amplicon that is frequently detected in small cell lung cancers. Am J Pathol 161: 207-216, 2002.
21. Pedersen N, Pedersen MW, Lan MS, Breslin MB and Poulsen HS: The insulinoma-associated 1: a novel promoter for targeted cancer gene therapy for small-cell lung cancer. Cancer Gene Ther: Epub ahead of print, 2005.
22. Pignon JP, Arriagada R and Ihde DC: A meta-analysis of thoracic radiotherapy for small-cell lung cancer. N Engl J Med 327: 1618-1624, 1992.
23. Warde P and Payne D: Does thoracic irradiation improve survival and local control in limited-stage small-cell carcinoma of the lung? A meta-analysis. J Clin Oncol 10: 890-895, 1992.
24. Pujol JL, Grenier J, Daures JP, Daver A, Pujol H and Michel FB: Serum fragment of cytokeratin subunit 19 measured by CYFRA 21-1 immunoradiometric assay as a marker of lung cancer. Cancer Res 53: 61-66, 1993.
25. Hennessy BT, Hanrahan EO and Daly PA: Non-Hodgkin lymphoma: an update. Lancet Oncol 5: 341-353, 2004.
26. Arlt A and Schafer H: NFkappaB-dependent chemoresistance in solid tumors. Int J Clin Pharmacol Ther 40: 336-347, 2002.